

SPECIFICATION

Electronic Version 1.2.8

Stylesheet Version 1.0

METHODS AND SYSTEMS FOR RAPID DEPLOYMENT OF A VALUATION SYSTEM

Cross Reference to Related Applications

This application is a continuation-in-part of U.S. patent application serial no. 09/741,211, filed December 19, 2000, which is incorporated herein by reference. The patent application serial no. 09/741,211, filed December 19, 2000, claims the benefit of U.S. Provisional Application No. 60/173,695, filed December 30, 1999, which is also hereby incorporated by reference in its entirety.

Background of Invention

[0001] This invention relates generally to valuation methods for financial instruments and more particularly to rapid valuation of large numbers of financial instruments.

[0002] A large number of assets such as loans, e.g., ten thousand loans or other financial instruments, sometimes become available for sale due to economic conditions, the planned or unplanned divestiture of assets or as the result of legal remedies. The sale of thousands of commercial loans or other financial instruments sometimes involving the equivalent of billions of dollars in assets must sometimes occur within a few months. Of course, the seller of assets wants to optimize the value of the portfolio, and will sometimes group the assets in "tranches." The term "tranche" as used herein is not limited to foreign notes but also includes assets and financial instrument groupings regardless of country or jurisdiction.

[0003] Bidders may submit bids on all tranches, or on only some tranches. In order to

win a tranche, a bidder typically must submit the highest bid for that tranche. In connection with determining a bid amount to submit on a particular tranche, a bidder often will engage underwriters to evaluate as many loans as possible within a tranche and within the available limited time. When the time for submitting a bid is about to expire, the bidder will evaluate the loans underwritten at that time, and then attempt to extrapolate a value to the loans that have not then been analyzed by the underwriters.

[0004] As a result of this process, a bidder may significantly undervalue a tranche and submit a bid that is not competitive or bid higher than the underwritten value and assume unquantified risk. Of course, since the objective is to win each tranche at a price that enables a bidder to earn a return, losing a tranche due to significant undervaluation of the tranche represents a lost opportunity.

[0005] Currently, business enterprises assess an acquisition or sale of assets and portfolios of assets on rapid schedules and often at great distances and varying time zones from the general management teams and functional heads which typically approve the offers for purchase or sale of these assets. Employees, partners and collaborators associated with the due diligence regarding the purchase of the assets are typically brought together for a relatively short duration of time to accomplish the due diligence. Typically due diligence activity is conducted in physical proximity to the sources of information associated with the assets. The collaborating personnel often do not have the benefit of training or knowledge of the complete set of analytical tools at their disposal nor do they have best practices from previous efforts of a similar nature.

[0006] Consolidation of employees and collaborators into a remote physical location for the duration of the due diligence effort is time consuming and expensive. In addition, persons on due diligence teams search for data and processes in an ad hoc fashion, typically relying on a small subset of other personnel who have detailed information about information sources, underwriting, analytical tools, reports, and completed analysis. The subset of individuals who have the information become bottlenecks within a due diligence time line, driving up due

diligence costs and adding time that could have otherwise been invested in more value added due diligence. Core information relevant to an asset portfolio bid is typically consolidated in a well controlled physical location, sometimes referred to as a war room.

[0007] It would be desirable to have a collaboration mechanism that brings the best of a company's previous experience and knowledge to bear on due diligence issues and that allows teams to become knowledgeable efficiently when considering assets purchases and to consolidate analytical building blocks in a repository that may be deployed quickly in future purchase deliberations, without the time and expense of the known war room approaches.

Summary of Invention

[0008] An integrated Internet based system, which organizes a company's experiences, operating procedures, best practices, information sources, competitive information and analytical tools on a server for easy storage and retrieval is disclosed. The invention is a method and a system to make available to a globally dispersed team, all of the real time analytical information and processes on a continuous basis over the worldwide web. In practice the goal when using the tool is to increase the profitability of an entity in a due diligence process while facilitating the ongoing operations for which the due diligence team members are responsible.

[0009] In an exemplary embodiment, invention is a method for collaborating on due diligence issues to affect efficient knowledge building within due diligence teams. The method includes the steps of accessing stored, accumulated knowledge in a repository from prior due diligence exercises, applying the accumulated knowledge to due diligence decisions criteria based on consolidated analytical building blocks of past due diligence exercises, and storing newly accumulated knowledge from the current due diligence exercise into the repository of accumulated knowledge for future use.

[0010]

In another exemplary embodiment, the invention is a system to implement the

process for managing due diligence. The system includes a computer, at least one server connected via network to the computer, the server having the connectivity to access global information over the internet. The system is configured to store accumulated knowledge in a repository from prior due diligence exercises, apply consolidated analytical building blocks of past due diligence exercises to due diligence decision criteria and store newly accumulated knowledge from the current due diligence exercise into the repository of accumulated knowledge.

[0011] In yet another exemplary embodiment, the invention is a computer program embodied on a computer readable medium. The computer program includes a code segment that sets up a directory structure to organize information into a centralized database and provides users access to a specific set of data stored in the centralized database to facilitate decision making in response to an inquiry. The computer program is capable of managing the information in the centralized database, managing project timeline with milestones, and providing the feedback to various participants in an easily readable set of graphical indicators to track project deliverables.

[0012] In yet a further embodiment, the invention is a centralized database which includes data corresponding to various projects, flow maps identifying sources and uses of the information, financial models and business process tools, best practices, valuation process and underwriting.

Brief Description of Drawings

[0013] Figure 1 is a flow diagram illustrating a known process for valuing a portfolio of loans.

[0014] Figure 2 is a flow diagram illustrating valuing a portfolio of loans in accordance with one embodiment of the present invention.

[0015] Figure 3 is a flow diagram illustrating, in more detail, one embodiment of a first portion of a rapid valuation process for large loan portfolios that breaks loans into categories of variance.

- [0016] Figure 4 is a flow diagram illustrating a second portion of a rapid valuation process for a large loan portfolios that aggregates from a basis to a tranche or portfolio basis.
- [0017] Figure 5 illustrates a probability distribution for exemplary assets whose recovery value is inferred.
- [0018] Figure 6 is a flow diagram of a supervised learning step of the process of Figure 3.
- [0019] Figure 7 is a flow diagram of an unsupervised learning step of the process of Figure 3.
- [0020] Figure 8 is an embodiment of the process for unsupervised learning.
- [0021] Figure 9 is an embodiment of the generation 1 (first pass) rapid loan valuation process.
- [0022] Figure 10 is a flow diagram of a fuzzy clustering method used in the unsupervised learning of Figure 8.
- [0023] Figure 11 is a pair of tables showing an example of model selection and model weighting for a rapid loan evaluation process.
- [0024] Figure 12 is a table showing exemplary attributes for a rapid loan valuation process.
- [0025] Figure 13 is a flow diagram of an exemplary clustering method for a rapid loan valuation process.
- [0026] Figure 14 is a system diagram; and Figure 15 is a diagram illustrating due diligence tools and processes.

Detailed Description

- [0027] Figure 1 is a diagram 10 illustrating a known process for valuing a large portfolio of assets 12 through an underwriting cycle and through to making a bid for purchasing asset portfolio 12, for example, in an auction. Figure 1 is a high

level overview of a typical underwriting and extrapolation process 10 which is not iterative and not automated. In diagram 10, underwriters underwrite 14 a number of individual assets from portfolio 12 to generate an underwritten first portion 16 and an untouched remainder portion 18. Before any of the assets are underwritten, first portion 16 is zero percent and remainder portion 18 is one hundred percent of portfolio 12. As the underwriting process progresses, first portion 16 increases and remainder portion 18 decreases. The objective is to underwrite as many assets as possible before a bid is submitted for the purchase of asset portfolio. The team of underwriters continues individually underwriting 14 until just before a bid must be submitted. A gross extrapolation 20 is made to evaluate remainder portion 18. The extrapolated value 20 becomes the non-underwritten inferred value 24. The gross extrapolation generates a valuation 24 for remainder portion 18. Valuation 22 is simply the total of the individual asset values in first portion 16. However, valuation 24 is a group valuation generated by extrapolation and may be discounted accordingly. Valuations 22 and 24 are then totaled to produce the portfolio asset value 26. Valuation processes are performed on each tranche of the portfolio.

[0028]

Figure 2 is a diagram illustrating one embodiment of a system 28 for rapid asset valuation. Included in Figure 2 are representations of process steps taken by system 28 in valuating asset portfolio 12. System 28 individually evaluates ("touches") every asset, except for a very small quantity 30 of untouched assets considered statistically insignificant or financially immaterial. Specifically, all assets in portfolio 12 other than quantity 30 undergo an iterative and adaptive valuation 32 in which the assets in portfolio 12 are individually valued, listed individually in tables and then selected from the tables and grouped into any desired or required groups or tranches for bidding purposes (as described below.) As in diagram 10, underwriters begin a full underwrite 14 of individual assets in portfolio 12 to produce a fully underwritten first portion 16 of assets. Underwriters also underwrite 34 a sample of assets in a second portion 36 of portfolio 12, and a computer 38 statistically infers 40 value for a third portion 42 of portfolio 12. Computer 38 also repetitively generates 44 tables (described below) showing

values assigned to the assets in portions 16, 36 and 42 as described below. In one embodiment, computer 38 is configured as a stand alone computer. In another embodiment, computer 38 is configured as a server connected to at least one client system through a network (shown and described in Figure 14), such as a wide-area network (WAN) or a local-area network (LAN).

[0029] For example, and still referring to Figure 2, an unsampled and non-underwritten portion 46 of a third portion 42 of portfolio 12 is subjected to a statistical inference procedure 40 using fuzzy-C means clustering ("FCM") and a composite High/Expected/ Low/ Timing/Risk ("HELTR") score to generate two categories 48 and 50. HELTR is defined as HHigh cash flow, EExpected cash flow, LLow cash flow, TTiming of cash flow (for example in months: 06, 718, 1936, 3760), and RRisk assessment of borrower (9boxer used by credit analysts). Category 48 is deemed to have sufficient commonality for evaluation as a whole. Category 50 is further divided into clusters 52 and 54 that are, in turn, further subdivided. Cluster 52 is divided into subclusters 56 and 58, while cluster 54 is subdivided into subclusters 60, 62 and 64. Cluster and subclusters are shown both in a "tree" chart 66 and as boxes in valuation block 68. These individual asset values are then regrouped into tranches 70, 72 and 74 for bid purposes. Any number of tranches could be assembled in any arrangement set by the seller.

[0030] Individual asset data (not shown) for each asset in portfolio 12 is entered into a database 76 from which selected data 78 is retrieved based on a given criteria 80 for the iterative and adaptive process 32. When criteria 80 is established for valuation of any asset, that established criteria 80 is stored in database 76 for use in valuating other asset data in database 76 which shares such an established criteria. Iterative and adaptive valuation process 32 thus develops 82 valuations (described below) and groups 84 them for use in bidding.

[0031] Figures 3 and 4 together form a flowchart 85 illustrating a functional overview of one embodiment of system 28 (shown in Figure 2) for evaluation of a large asset portfolio 12. Valuation procedures 14, 34 and 40 (see also Figure 2) are simultaneously and sequentially used in system 28 in a manner described below.

As described above, full underwriting 14 is a first type of valuation procedure. Grouping and sampling underwriting 34 with full underwriting of the samples is a second type of valuation procedure. Statistical inference 40 is a third type of valuation procedure, which is an automated grouping and automated valuation. Procedures 14, 34 and 40 are based on objective criteria established as described below.

[0032] "Underwriting" as used herein means a process in which a person ("underwriter") reviews an asset in accordance with established principles and determines a current purchase price for buying the asset. During underwriting, the underwriter uses pre-existing or established criteria 80 for the valuations. "Criteria" means rules relevant to asset value and a rating based on such categories. For example, as a criteria, an underwriter might determine three years of cash flow history of the borrower to be a category of information relevant to asset valuation and might give a certain rating to various levels of cash flow.

[0033] Full underwriting 14 is done in two ways, a full cash basis manner 86 and a partial cash basis manner 88. Both full cash basis manner 86 and partial cash basis manner 88 start with sets 90 and 92 of assets that are fully individually reviewed 14 (see Figure 2). Such full review 14 is usually due to the large dollar, or other appropriate currency, amounts of the assets being reviewed relative to other assets in the portfolio or due to the borrower being so well known or so reliable that the assets can be quickly and reliably fully underwritten or the assets are marked to market such that there is very little variance associated with the value of said assets. Asset set 90 is evaluated by underwriters 94 and each asset in set 90 receives a valuation with very little variation such as an asset backed with cash or a tradable commodity with full cash value and is placed in a full value table 96. Selected individual values for assets in table 96 are stored as a fully underwritten group value 98.

[0034] Set 92 is evaluated by a team of underwriters 100, which could be the same as team 94, but each asset receives a discounted or partial value and is placed in a partial value table 102. Selected individual values for assets in a tranche in table

102 are stored as a partial value fully underwritten group value 104. Criteria 80 (shown in Figure 2) for full cash basis manner 86 and partial cash basis manner 88 are stored in database 76 (shown in Figure 2) in a digital storage memory (not shown) of computer 38 (shown in Figure 2) for use in supervised learning 206 and unsupervised learning 208 of automated valuation 40.

[0035] Sampling underwriting 34 is accomplished using two procedures, a full sampling 106 procedure and a partial sampling 108 procedure. Full sampling 106 is utilized for categories of large assets and includes a one hundred percent sampling 110 of the sample groups in the categories of assets being sampled. The assets in full sampling 106 are not individually underwritten but rather are underwritten in full sampling groups 112 based on a determined commonality. A resulting full sampling group valuation (not shown) is created and then desegregated based on a rule 114 to generate an individual full sample asset value table 116. Individual full sample asset values in table 116 are then uploaded electronically into any full sampling group valuation 118 required for bidding as suggested by the grouping of assets in a tranche. The number of assets in an underwriting sample grouping can be as little as one to any number of assets. Partial sampling 108 is for medium categories of assets and includes forming a cluster sample group 120 by one hundred percent sampling of a representative group from within a cluster of the groups being sampled and random sampling of the other groups in the cluster. In partial sampling 108, all groups are sampled, but some are partly valued by extrapolation from cluster sample group 120. Partial sampling 108 includes an asset level re-underwrite 122 with manual data entry 125 to produce an alpha credit analyst table 126 which is given an asset class adjustment 128 to produce an adjusted credit analyst table 130. As described above, individual assets are selected from adjusted credit analyst table 130 according to tranche grouping to produce a partial sampling credit value 132 for use in bidding on tranche 70 (shown in Figure 2).

[0036] Automatic valuation procedure 40 utilizes supervised learning process 206, an unsupervised learning process 208 and an upload from a statistical inferencing algorithm 134 to generate an underwriting clusters table 136 which is stored in a

digital storage device. In supervised learning process 206, an experienced underwriter who knows what questions to ask to establish value, assists the computer in determining whether or not an asset is a good investment and how to value the asset. In unsupervised learning process 208, the computer segments and classifies assets and objectively self-evaluates the assets based on feedback from the data. An underwriter periodically reviews the unsupervised learning process 208 to determine whether the computer is making sensible underwriting conclusions. The computer uses statistical algorithms 134 to make its inferences. For example, but not by way of limitation, one embodiment uses the Design For Six Sigma ("DFSS") quality paradigm developed and used by General Electric Company and applied in a Due Diligence ("DD") asset valuation process using a multi-generational product development ("MGPD") mode to value the asset data with increasing accuracy. Learning processes 206 and 208 incorporate the accumulated knowledge as the valuation progresses into cash flow recovery and probability of recovery calculations on an ongoing, real time basis. Supervised learning process 206 uses business rules to identify clusters of assets having common aspects for valuation purposes. Unsupervised learning process 208 uses feedback from prior data valuations performed by procedure 40 to determine if progress is being made with respect to increasing valuation confidence. Identification of all available raw data and discovery of interrelationships of clusters of these available raw data is possible due to the use of high-speed computers, as is described below.

[0037] In one exemplary embodiment, a fuzzy clustering means ("FCM") process of unsupervised organization of raw data using a HELTR scoring technique is employed to infer valuations of credit scores onto assets in portfolios, as described below. Such clustering techniques have been developed in response to more sophisticated classification segments to describe assets and high asset counts in portfolios that must be assessed in time periods that do not allow manual processing.

[0038] One exemplary method first organizes valuation scores (static and/or probabilistic recoveries) in a computerized system. Adjustments are then made to the valuation scores for special factors and business decisions. Then a

reconciliation of multiple valuation scores describing the same asset and an overall adjustment to interview/override the inferred valuation is performed.

[0039] Organizing valuation scores is performed by collating, in electronic form, a cluster number, a cluster name, descriptive attributes of the cluster(s), probabilistic recovery values (an illustrative example is a HELTR score) and the underwriter's confidence in each cluster's valuation based upon the strengths of each cluster's descriptive attributes. The cluster number is a unique identifier of a specific set of descriptive attributes that are facts about an asset which a person skilled in evaluations uses to assess value of an asset. Examples of descriptive attributes include, but are not limited to, payment status, asset type, borrower's credit worthiness expressed as a score, location and seniority of a claim. The cluster name is, in one embodiment, an alpha-numeric name that describes the cluster's descriptive attributes or sources. One example of descriptive attributes is found in Figure 12, described below.

[0040] Descriptive attributes are the facts or dimensions or vectors that were used to develop the asset's value. Computer logic is used to check for replicated clusters, if any, and alert the analysts or underwriters.

[0041] Because each asset can be described by many combinations of descriptive attributes, various levels of value for the same asset may occur. Probabilistic recovery values or credit score or any numerical indication of the asset's worth are indicators of worth designated at the discrete asset level. All of the information from the various descriptive attributes is synthesized such that a purchase or sale price can be ascertained as a fixed value or a probabilistic one. An illustrative embodiment used herein is the HELTR score. Each cluster has a unique set of descriptive attributes and designated HELTR score.

[0042] Every cluster's unique attributes contribute to a valuation of cluster value. Different combinations of attributes provide a higher confidence or confidence interval of a particular cluster's score. For example, if any asset was described as a green piece of paper with height equal to 2.5 and width equal to 5 one might ascribe a value of 0 to 1000 dollars and place very little confidence in this

assessment. If this same asset was described with one more fact or attribute or vector as being a real \$20 US bill, one would place a very high confidence factor on this cluster value of \$20 US dollars.

[0043] A cluster's valuation and confidence is determined at a point in time and recorded. Sometimes new information becomes available and the analyst would like to alter the value(s). The value is altered manually or automatically with a data field and decision rules, in the automated fashion via computer code. The prior values are manipulated to reflect new information. As an illustrative example, assume the prior cluster confidence was recorded at 0.1 and it is learned that a different asset with exact descriptive attributes as in this cluster just sold for over the predicted most probable value. Rules were in effect such that if this event occurred, cluster confidence is multiplied by 10. $0.1 \times 10 = 1$ which is the revised cluster confidence.

[0044]

The purpose of such a process is to reconcile multiple scores for the same asset, controlling for the confidence associated with each source of valuation of each dimension of valuation. Using the HELTR as an illustrative example with sample data points on a particular asset:

[t3]

Cluster Number	Cluster Name	High	Low	Time	Value	High	Low	Time
1	Low persons receive	15	15	1	1	151.05W(15)	151.05W(15)	151.05W(15)
2	Asset classification include	15	15	1	1	151.05W(15)	151.05W(15)	151.05W(15)
3	Coordinate use -	15	15	1	1	151.05W(15)	151.05W(15)	151.05W(15)
4	Inventory	15	15	1	1	151.05W(15)	151.05W(15)	151.05W(15)
5		15	15	1	1	151.05W(15)	151.05W(15)	151.05W(15)

[0045]

The cluster consensus valuation is a high value of .6999, most likely .4792, low .2374 with a timing of 2.6059. Different logic can be applied to manipulate any

of the weights.

[0046] The consensus scores are developed in the context of global assumptions. Should a global assumption change occur, process steps 128, 138 are included in the methodology to weight the consensus score. Illustrative examples are fraud discovery in certain valuation factors, macroeconomic changes, fungible market value established for an asset class, and loss of or increase of inferred asset valuation methodologies relative to other methodologies being employed.

[0047] In another embodiment, a cross correlation tool is used to quickly understand and describe the composition of a portfolio. Typically, the tool is used to correlate a response of a user selected variable versus other variables in an asset portfolio. The tool quickly identifies unexpectedly high or low correlation between two attribute variables and the response variable. Attribute variables are of two types, continuous and categorical. The cross correlations are computed by the correlation tool between all variables of interest and their bin or level and presented, in one embodiment, in a two dimensional matrix for easy identification of trends amongst the assets in the portfolios.

[0048] First, the cross-correlation tool identifies attribute variables in the portfolio of assets as one of continuous or categorical. For each variable aggregation levels are computed by bins for continuous variables and by value for categorical variables.

[0049] A user looking to identify correlations with the tool will select a response variable, Y_r , for example, an expected recovery or count. For all combinations of pairs of attribute variables (x_1 and x_2) and their levels (a and b), compute the average value of the response variable, Y_r , according to:

[0050]
$$Y_r = \text{sum}(Y(x_1 = a \text{ and } x_2 = b)) / \text{count}(x_1 = a \text{ and } x_2 = b).$$

[0051] An expected value, Y_{expect} , of the response variable is calculated according to:

[0052]
$$Y_{\text{expect}} = (\text{sum}(Y(x_1=a)) * \text{count}(x_1=a) + \text{sum}(Y(x_2=b)) * \text{count}(x_2=b)) / (\text{count}(x_1=a) * \text{count}(x_2=b)).$$

[0053] A deviation, Y_{error} , of the chosen response variable, Y_r , from the expected value, Y_{expect} , using weighted values of occurrence of $x_1=a$ and $x_2=b$ separately, is calculated by:

[0054]
$$Y_{\text{error}} = Y_r - Y_{\text{expect}}.$$

[0055] In one embodiment, expected values and deviations are displayed in multi-dimensional displays to make variations from expected values easy to identify.

[0056] In another exemplary embodiment, a transfer function process that converts raw data into the ultimate bid price is used, as described below. Table 136 is electronically adjusted using modified coefficients developed in procedures 14, 34 and 40 to give a coefficient adjustment to a credit score 138 for the asset and to generate an adjusted credit analyst table 140 of inferred individual asset credit values. Individual asset values are taken from table 140 as required by tranche grouping to generate an inferred credit valuation 142. Finally an extrapolation is made on the negligible remainder 30 of "untouched" assets to generate a table of untouched assets 144. Values from table 144 are selected to generate an untouched asset valuation.

[0057] Full cash valuation 98, partial cash valuation 104, full sampling credit valuation 118, partial credit values 132, inferred credit value 142 and any value assigned from untouched asset table 144 are cumulated and are mutually exclusive with the priority being full cash valuation 98 to inferred credit value 142 consecutively. A sum of the valuations represents value of the portfolio.

[0058] Figure 4 is a flow diagram of a bid preparation stage 168 performed by system 28 (shown in Figure 2). The cumulated valuations 98, 104, 118, 132, 142 and 144 are combined in a risk preference loan level valuation step 146. A deterministic cash flow bridge 148 is produced using a cash flow timing table 150 to develop a stochastic cash flow bridge 152. A stochastic or probabilistic cash flow bridge 152 is created and used to determine a proposed tranche bid price 154 to which is applied a tranche model 156 iteratively until a certain threshold 158 is reached. Threshold 158 is, for example, an internal rate of return ("IRR") greater than some

value, a certain time to profit ("TTP"), and a positive net present value ("NPV").

[0059] In general, NPV is defined as:

$$NPV = C_0 + \frac{C_1}{1+r} \quad (\text{Equation A})$$

where C_0 is the investment at time 0, C_1 is the expected pay-off at time 1, and r is the discount factor. The basic idea is that a dollar today is worth more than a dollar tomorrow.

In the case of insurance policies, NPV is defined as:

$$NPV = \sum P - \sum E - \left(\sum C \right) \cdot \frac{A}{E_r} \quad (\text{Equation B})$$

[0060] where P is the premium, E is the expected nominal cost, and C is the claim cost. In essence, Equation B is how net income as the difference of profit and weighted expected risk is generated. Note that the summation is summing across all the policies in a specific segment. Also note that all the premium, nominal cost, and claim cost have been discounted before entering the equation. As a result, a profitability score is generated.

[0061] If threshold conditions 160 are met, bid 154 is subjected to a simulated bid opening analysis 161 to predict whether the bid can be expected to be a winning bid. An outcome of a sealed bid auction depends on sizes of the bids received from each bidder. Execution of the auction involves opening all of the bids and selling the items up for auction to the highest bidder. In traditional sealed bid auctions, bidders are not allowed to change their bids once their bid is submitted and bidders do not know the bids placed by other bidders until the bids are opened, making the outcome of the auction uncertain. By placing higher bids, a probability that the auction will be won is higher, but value gain is lower if it was possible to have won the auction at a lower price.

[0062] Simulating competitive bidding increases the probability of capturing the highest upside of profitability by setting a range of bid/sale prices that have a propensity to exhaust any competing bidder's purses before one's own purse such that the most desirable assets transact with the highest preservation of capital. Pricing decisions are brought into focus by an analytically robust process because pure anecdotal business judgment can be augmented by a data driven approach not subject to a hidden agenda, personality or unilateral knowledge.

[0063] Each potential bidder has a range of possible bids that might be submitted to a sealed bid auction. The range of bids can be expressed as a statistical distribution. By stochastically sampling from a distribution of bid values, one possible auction scenario may be simulated. Further by using an iterative sampling technique, for example a Monte Carlo analysis, many scenarios are simulated to produce a distribution of outcomes. The distribution of outcomes include a probability of winning the auction item(s) and the value gain. By varying the value of one's own

bid, a probability of winning the auction against ones own bid price can be determined.

[0064] The following core elements are used to simulate a competitive bidding yield, codification of market rules and contracts into computerized business rules, codification of potential competition/market forces, forecasted budgets and priorities into a preference matrix, one's own bidding capacity, preferences, risk/return tradeoffs agreed to codified into a preference matrix, and a computerized stochastic optimization.

[0065] Analysis 160 simulates a competitive environment with other companies having various financial capabilities bidding against the bids calculated by system 28. In one embodiment, analysis 160, for example and without limitation, includes a total bid limit such as would be the case where the total value of the assets exceed the financial capabilities of the entity using system 28. In one embodiment, analysis 160 might assess the profitability, in such case of limited resources to bid, of bidding on various combinations of tranches. Analysis 160 also takes into account past history in bidding against known competitors and information on the various types of assets preferred by competing bidders. In analysis 160, the tranche bid is then evaluated and set by management 162 and a final tranche bid 164 made. All valuations prior to the making of the bid 164 can be repeated as desired. Further, since the process is self-adjusting and iterative, the tranche bid price 164 tends to climb upward with each iteration as more and more value is found by the iterations performed by system 28.

[0066] The process described by flowchart 85 includes an evaluation stage 166 (shown in Figure 3) and a bid preparation stage 168 (shown in Figure 4). Evaluation stage 166 includes procedures 14, 34 and 40. Evaluation stage 166 runs constantly until stopped, with the automatic valuation procedure 40 and sampling procedures 34 attempting to find extra value in various assets or categories of assets.

[0067] Referring once again to Figure 2, and in accordance with rapid asset valuation, data categories 170, 172 and 174 within the assets of portfolio 12 are identified

on each asset and stored in database 76. Iterative and adaptive valuation process 32 takes portions of selected data 78 and applies criteria 80 to the portions of selected data 78 in a statistical manner to increase the known asset value rather than the asset value being a gross extrapolation 20. In accordance with method 28 the assets are divided into at least first portion 16, second portion 36 and third portion or remainder 42. Using procedure 14, the assets in portion 16 are fully underwritten to determine valuation 98 and partial value fully underwritten valuation 104 and to establish criteria 80 for such valuation. Using procedure 34, process 28 samples a quantity of assets from second portion 36 representative of groups in second portion 36 to determine full sampling group valuation 118 and partial sampling credit values 132 for second portion 36 and to establish additional criteria 80 for such valuation. Using procedure 40, partially supervised learning process 206 and partially unsupervised learning process 208 are performed by an automated analyzer such as computer 38 of Figure 2. In order to learn, the automated analyzer extracts established criteria 80 and selected data 78 as to third portion or remainder 42 and divides third portion 42 into portions 46, and then further divides each portion 46 into categories 48 and 50 and category 50 into clusters 52, 54 and clusters 52, 54 into subclusters 56, 58, 60, 62 and 64 using criteria 80 imported from database 76 and each of processes 206 and 208. Individual asset valuations are established for the assets in subclusters 56, 58, 60, 62 and 64 by statistical inference.

[0068]

The individual asset valuations are listed in cluster tables 136 (see Figure 3) and after adjustment 138, listed in a credit analyst table 140. The established criteria 80 are objective since criteria 80 come from database 76 where they have been placed during full underwriting procedure 14 and sample underwriting procedure 34. In other words, information obtained in full value table 96, partial value table 102, table 116, alpha credit analyst table 126, adjusted credit analyst table 130, adjusted credit analyst table 140 and untouched asset table 144 for all assets is placed into database 76 in a digital storage device, such as the hard disk storage 178 of computer 38, and correlations are made by procedure 40 with criteria 80 from procedures 14 and 34. During procedure 40, criteria 80 which are

of statistical significance with an acceptable degree of reliability, are entered. That is, procedure 40 iteratively learns as it values and establishes criteria 80.

Supervised learning process 206 and unsupervised learning process 208 increase the accuracy of statistically inferred valuation 142 by correlating to established criteria 80 in database 76 on assets in fully underwritten first portion 16 and assets in sample underwritten second portion 36. Selected data 78 related to one or more assets in third portion 42 similar to selected data 78 on assets in portions 16 and/or 36 are located in database 76 and then by statistical inference, a value for each asset in third portion 42 is determined from the located information.

[0069] During the process described by flowchart 85, assets are valued at an individual asset level, and the individual asset values are tabulated or grouped in one or more combinations. To have maximum flexibility for various bidding scenarios, any subset of portfolio 12 is valued and priced separately in a particular time frame. In known process 10, if a seller of assets regroups the assets, for example from groupings by asset company to groupings by geographical location of borrowers, revaluation of bids may be inadequate because gross extrapolation 20 will need to be performed. In using system 28, because individual asset values are developed and listed in tables 96, 102, 116, 130, 140 and 144, these values can be electronically regrouped into different valuations 98, 104, 118, 132, 142 whose food chain selection criteria is mutually exclusive and selectable by the analysts conducting the evaluation and is further described below. If the seller groups the assets, then grouping according to seller groups or tranches is easily made and an appropriate valuation 146 developed for that tranche. The individual asset values are thus easily regrouped for third portion 42 to objectively obtain an inferred valuation 142 for that group or tranche.

[0070] Many methods may be employed to establish asset value. Depending upon the objectives of the valuation, the relative merits of different valuation methodologies establish the desirability of the valuation techniques for a particular asset. One methodology is similar to a food chain which preserves assumption development methods yet selects the intervals with the highest confidence intervals.

[0071] In one introductory illustrative example of a food chain, one may prefer to value a financial asset more by what similar assets trade in the open market for versus an individual's opinion. In rank order, the market-to-market value is selected over an individual's opinion.

[0072] In the same way assets in a portfolio with a forecasted cash flow recovery may be evaluated by a number of valuation techniques. The typical objective is to establish, with as high a probability available, what the future cash flow will be. The valuation methodologies are ranked in order of their capability to accurately quantify cash flow, or cash equivalent, forecasts with the least downside variances and/or maximum upside variances. The asset is valued by all available methods that have merit, or may have business logic rules to eliminate duplicate work when it is known that more accurate methods will preclude the need to assess an asset's valuation once the best method has been employed.

[0073] In order to provide the best forecast of asset value, assets are evaluated by each method within a food chain until such time as they are valued by the best available method for each particular asset. Once this best value is found, the asset is said to have its value, irrespective to other values lower (with more variance) in the food chain and is sent to the completed state.

[0074] As an example, a portfolio of assets is evaluated using a food chain. The first valuation method in the food chain is the one which most closely matches the valuation objectives namely to find the value with the highest degree of accuracy (tightest confidence interval). As soon as the asset is valued by a methodology for which a value was established for that unique asset, it is sent to the valuation table and removed from any further steps in the food chain. A list of assets from the original portfolio that did not match any valuation methods is kept in the untouched asset table. The objective is to drive this untouched table to zero assets.

[0075] One example of a food chain is as follows, in order of preference: (a) 100% cash in hand for the asset, (b) partial cash in hand for the asset, (c) liquid market value for like asset, (d) direct underwrite, and (e) inferred underwrite.[0016] The

food chain approach provides an ability to find the best probability distribution shape, reduces probability distribution variance (especially on the downside tails), provides capability to establish probability distributions quickly while preserving all available knowledge in the constituencies and provides the ability to provide the best estimate of value at any point in the discovery process.

[0076] As shown in Figure 4, the general framework of bid preparation stage 168 is to price bid 164 similar to option valuation paradigms where the winning investor will have the right, but not the obligation, to recover the investment. The values are desegregated into three parts for each tranche, a time value of money component, an inherent value component and a probable cash flow component. The time value of money and the inherent value are deterministically calculated and have little variation once established. The time value of money is computed by taking a firm's cost of capital for a low risk investment multiplied by the investment for the applicable period which represents an opportunity for alternate investment that is foregone in order to make the present investment. Inherent value is a known liquid asset value, which is in excess of the purchase price and is available immediately after taking control of the assets. One embodiment is a well traded security purchased below market value as part of a portfolio. Probable cash flow variance is a function of the assumptions a due diligence team makes and the process it selects to convert raw data into a cash flow recovery stream. The systems described herein are configured to reduce negative variances and find value.

[0077] Figure 5 is a triangular probability distribution graph for a typical minimum three-point asset evaluation 180. In accordance with process 40 a minimum of three cases per financial instrument are evaluated. A vertical axis 182 denotes increasing probability and a horizontal axis 184 denotes increasing portion of recovery. A liquidation or worst case percentage 186 of a face value line 188, a best case percentage 190 of face value 188, and a most probable case percentage and recovery value 192 of face value 188 are shown. The probability of worse case percentage 186 is zero, the probability of best case scenario 190 is zero and a probability 194 of the most probable percentage 192 of recovery is a value represented by point 196. The size of an area 198 under a curve 200 defined by a

line connecting points 186, 196 and 190 is representative of value in the asset. The notational asset value holds to an area 202 of a rectangle bounded by a 100 % probability line 204 of a 100 % recovery of face value 188 is a measure of the portion of face value 188 that can be attributed to the asset represented by curve 200. Points 186, 196 and 190 and lines 188 and 204, and thus areas 198 and 202, will vary depending on selected data 78 chosen for the asset in question and criteria 80 applied to the asset and ascribed probabilities of asset value recovery. Horizontal axis 184 can be expressed in currency units (e.g. dollars) rather than percentage of face value. When currency units are used, areas 198 under curves 200 for different assets will be in currency units and thus areas 198 relate to each other in magnitude and hence in significance to overall bids 70, 72 and 74. The more that is known about the asset, the more curve 200 can be refined. Statistics are applied to curve 200 as criteria 80 are established to help establish the location of points 186, 196 and 190 and hence area 198 and thus the expected value of the asset. The timing of cash flows, which affects value, can be based upon histogram results of the timing attributes.

[0078] For example, the cash flow recovery timing can be broken down into three bins of 0–6 months, 7–12 months, 13–18 months, and so on. The automated analyzer 38 using algorithm 134 can select the bin width based upon a sensitivity study trade off of timing to valuation against the gauge recovery and rate determined possible by an underwriter. In an exemplary embodiment, a minimum of 4 bins should be utilized when the discount factor is more than 25%. For a discount factor between 10 and 25, a minimum of 6 bins should be used to cover the likely recovery periods.

[0079] In accordance with procedure 40 other sources of data are chosen that an underwriter would be able to utilize to assess value in a financial instrument. Criteria 80, established by underwriting teams 94, 100 114, 122 and 140 in procedures 14 and 34, are useful in that regard. In accordance with the process described by flowchart 85, raw data is turned into a recovery and a rule set is selected to apply a valuation to the raw data and this rule set is coded into the valuation database in the form of criteria 80. Each time a cluster is touched by

multiple hits during a valuation in procedures 14, 34 or 40, a consensus forecast is developed and applied to the cluster. In accordance with system 28, the probability distributions of cash flows and timing at the tranche level is determined by developing valuation transfer function 146 at the asset level which will take raw data, rationalize the assumptions that data will generate and aggregate the valuations of the individual assets in the tranche.

[0080] Since all recoveries are not homogeneous, a method to establish the variability of cash flow recoveries is provided. Individual assets are clustered by group exposure. As much face value as possible is traditionally underwritten in the time permitted, recognizing that a sizable sample remains for clustering. Clustering reserves are estimated using a sample size equal to one hundred forty five plus 2.65 % of the face count and a regression analysis of variance. This produces sample sizes of thirty for a face count of 100 assets, 150 for a face count of 1,000 assets, 400 for a face count of 5,000 assets, 500 for a face count of 10,000 assets, and 600 for a face count of 20,000 assets.

[0081] During statistical inference procedure 40, assets remaining in third portion 42 of portfolio 12 are clustered by descriptive underwriting attributes or criteria 80 and random samples are taken from each cluster and the sample underwritten. In one embodiment, sampling from a cluster in procedure 40 is stopped when asset level mean variance falls below 10%. In another embodiment, sampling is stopped when tranche level mean variance falls below 15%. Portfolio mean variance is not used as a stop point if the potential unit of sale is less than the entire portfolio. In accordance with procedure 40, recovery valuation of the cluster sampling is inferred onto the corresponding cluster population. In using system 28, the goal is to touch each inferred asset valuation via three or more unique clusters. During procedure 40 a cluster's underwriting confidence and descriptive attribute's relevance is weighed.

[0082]

By way of example, without limitation, 0 = no confidence that this cluster's descriptive attributes will provide a meaningful valuation; 1 = complete confidence that this cluster's descriptive attributes will provide as accurate of a valuation as

Table A Portfolio attributes

Borrower Size (by Borrower Group UPB)
Secured
Syndicated (yes / no)
Guaranteed
Loan Type (Term, Revolving, etc.)
% UPB from Loans in First Position
Collection Score (0=Bad 1=Good)
12-month collections % of UPB
% of Last Payment for Principal
Borrower Loans
Loan's portion of borrower UPB
Single Family Residence
Residential
Retail
Industrial
Hospital
Hospitality
Multifamily
Land Developed / Undeveloped / Other
Office
Stock / Mergers Loans

[0084]

Segmentation of the asset attributes is accomplished by encoding of attributes into "dummy variables". For example, a common asset attribute is "Has borrower

made a payment in the last 12 months?", which would be encoded in a variable as a "1" if the answer is yes, and "0" otherwise. Similar "dummy variables" are used for other asset attributes.

[0085] The segmentation procedure is completed by using any statistical procedure which process the encoded asset attributes in such a way so as to segment the portfolio into groups of similar assets. One such algorithm is K-means clustering. In an example, where three asset attributes, Unpaid Principal Balance (UPB), Probability of Payment, a scale from 0 to 1; and Secured Score, a probability of being secured by real estate collateral are used, the assets might be classified into five groups with similar attributes.

[0086]

Once the groupings of assets is made, the number of samples to be taken and submitted for further underwriting review is calculated by establishing the confidence level with which statements can be made about the total recoveries in each segment (k), establishing the precision with which one wishes to estimate the total recoveries in each segment (h) and providing an *a priori* estimate of the level and range of recoveries as a percentage of total Unpaid Principal Balance (UPB) (R), according to:

$$Var(\hat{Y}_k) = n \left[1 - \frac{n}{N} \right] \times \frac{\left[\sum_{i=1}^N \lambda_i \right]^2}{\left[\sum_{i=1}^N x_i \right]^2} \times \frac{\sum_{i=1}^N (y_i - Rx_i)^2}{N-1}$$

n – sample size

N – cluster size

x_i – UPB for sample i

y_i – recovery for sample i

$$R = \frac{\sum_{i=1}^N y_i}{\sum_{i=1}^N x_i} = \text{cluster expected recovery \%}$$

$$h^2 = k^2 \times n \left[1 - \frac{n}{N} \right] \times \frac{\left[\sum_{i=1}^N x_i \right]^2}{\left[\sum_{i=1}^N x_i \right]^2} \times \frac{\sum_{i=1}^N (y_i - Rx_i)^2}{N-1} \quad (\text{Equation C})$$

h – error tolerance for estimating $Y = \sum_{i=1}^N y_i$ with \hat{Y}_k

$$\hat{Y}_k - \hat{R} \times \sum_{i=1}^N x_i = \frac{\sum_{i=1}^N y_i}{\sum_{i=1}^N \lambda_i} \times \sum_{i=1}^N x_i - \frac{\sum_{i=1}^N y_i \lambda_i}{\sum_{i=1}^N \lambda_i} \times \sum_{i=1}^N x_i \quad (\text{Equation D})$$

k – constant in Tchebychev's Formula

$$|\hat{Y}_k - \mu_{Y_k}| \leq k \sqrt{Var(\hat{Y}_k)} \text{ with probability } \geq 1 - \frac{1}{k^2}$$

By solving Equation C for n , required sample size for the given cluster is obtained. Solving Equation C further allows the user to state, with probability $1 - \frac{1}{k^2}$ the calculated sample size, n , and associated underwritten values will estimate the total cluster recoveries to within an error of h , assuming that estimates of total segment recoveries are determined using Equation D.

[0087]

In practice, it is difficult to estimate variability in total recoveries without available data. A spreadsheet tool implements the above by generating data in a

recovery is valued with continuous probabilities at the unit of sale, which in one embodiment is a tranche. In the use of system 28, internal rate of return ("IRR") and variance would then be assessed. Preferred tranches have lower variances for a given IRR. The probability of each tranche's net present value ("NPV") to be above 0 is assessed using the project's discount rate. A discount rate is determined from the opportunity cost of capital, plus FX swap cost, plus risks in general uncertainties inherent in the variances of forecasted cash flow recovery. If it appears that there is more than a five-percent certainty that the project will have a negative NPV, no bid is made. Deal evaluation is by tranche with decision criteria being IRR, risk variance of the IRR in a tranche, estimated willingness and ability of the tranche to pay, time to profit ("TPP") and the risk variance in the payback by tranche, and NPV of the expected cash flow by tranche discounted to risk free rate.

[0090] In competitive bid circumstances when the content of asset portfolios is not negotiable, the investor or seller has a strong financial incentive to select only the portions of total assets available for transaction that will give their aggregated financial structure the best risk/return. Meeting minimum risk/return expected values with assets that will have a higher probability of maximum upside probabilities is even more attractive to investors.

[0091] The aggregated portfolio is divided into separately marketable sub portfolios or tranches. Each tranche has a forecasted cash flow probability distribution and time duration from prior analytics. These tranches are then given a trial price. The new assets are combined with the existing asset performance of the selling or buying party and subjected to Monte Carlo case generation (with associated cross correlations accounted for).

[0092] The tranche selection process includes a random selection of tranches not to buy. Once the portfolio effects take on a pattern, the best selection of tranches to purchase, at what price, subject to constraints is found by stochastic optimization.

[0093] Using NPV can be misleading due to the effects associated with double discounting which will occur when pessimistic case scenarios are discounted to obtain PV. Using time to profit is used to overcome this limitation and the marginal

capital cost or risk free rate is used in the discounting as determined by analysts conducting the evaluation.

[0094] Supervised learning process 206 of inferred valuation procedure 40 and steps 120, 122 and 126 of partial sampling procedure 108 have substantial similarity in that the underwriter is actively involved in the process, but the process is automated. Figure 6 is a flow diagram illustrating a process 210 for automated underwriting of segmentable financial instrument assets. First clusters of financial instruments are defined 212 by common attributes. An expert opinion 214 of value is given for selected samples from the defined clusters based upon the attributes. This opinion is used in a sample underwriting process 216 and values are checked for combinations of attributes and reconciled 218. Process 210 then selects and sets 220 the individual attributes to be used and then classifies 222 individual assets into clusters. Cluster valuation is applied 224 to each cluster asset. Using the cluster valuation, the values are desegregated by a rule 226 to create a credit analyst table 228.

[0095] Figure 7 is a flow diagram of one exemplary embodiment of unsupervised learning 208 that includes several modules. A data acquisition module 230 collects relevant data 78 wherever available. A variable selection module 232 identifies the asset relevant variables deemed critical by credit review or with the most discriminate power in separating various asset groups. A hierarchical segmentation module 234 segments the entire portfolio of assets into bins based on critical variables selected by analysts. A FCM module 236 further classifies each bin into clusters based on natural structure of the asset data. An underwriting review module 238 assigns projected cash flow and risk scores 138 (shown in Figure 3) to each cluster. This score is then supplied to the individual asset values in credit analyst table 136 for the assets from the clusters being adjusted in procedure 40 to produce adjusted credit analyst table 140. The process is iterative and continuous and can be performed by computer so that it continues while standard underwriting is being performed elsewhere.

[0096] Figure 8 illustrates an alternate exemplary inferred valuation process 240 used

in place of the process described in Figures 3 and 4. In alternate process 240, a seven-step process is used to rapidly value a real estate loan portfolio using a combination of full underwriting, partial underwriting and inferred valuation. First, assets are sampled 242 according to risk. Second, assets are underwritten 244, and valuations recorded. Third, market value clusters are formed 246, such as by FCM, as described below. Fourth, regression models are built 248, for the underwritten assets. A best model is selected 250, for the underwritten assets from among those built 248 earlier. Sixth, the counts for the selected models are calculated 252. Seventh, models are applied 254, as selected 250 to non-underwritten or inferentially valued portion 42 of portfolio 12 in a manner weighted by the counts to predict individual values for each of the non-underwritten assets. The individual asset values produced according to process 240 are then placed in adjusted credit analyst table 140 (see Figure 3).

[0097] In sampling assets 242, underwriters use stratified random sampling to select assets for detailed review. Strata are constructed from collateral attributes. Examples of collateral attributes for real estate portfolios include, collateral usage (commercial or residential), previous appraisal amount, market value cluster (predicted from previous appraisal amount, land area, building area, current appraisal amount, court auction realized price, property type and property location. Typically, assets are sampled in an adverse manner, i.e., purposely selected from a list ordered by decreasing Unpaid Principal Balance ("UPB") or Previous Appraisal Amount ("PAA").

[0098] Underwriting 244 is a largely manual process in which expert underwriters ascribe a notion of worth to collateral assets. The underwritten valuations are stored in a master database table, such as database 76 (shown in Figure 2). Valuations are typically summarized in terms of monetary units (e.g., 100,000 KRW), at then current market prices.

[0099] Figure 9 is a high level overview 290 of the automated portion of the process employed by system 28. Automated procedures are used by underwriters to assist in full underwriting based on procedure 34 (see also Figure 3). Knowledge captured

in procedure 34 is applied in inferred valuation procedure 40 to reduce cost and uncertainty in due diligence valuations of financial instruments and to reduce cost and variability between due diligence valuations. The valuations are subjected to a cash flow model which includes asset level valuation 146, deterministic cash flow bridge 148, stochastic cash flow bridge 152 and cash flow table 150. The resultant bid valuation 154 is subjected to gaming strategies 160 and management adjustments 162 to produce the final bid 164.

[0100] Figure 10 is a flow diagram of an exemplary embodiment of forming clusters 246. In forming clusters 246, underwriters, with the aid of algorithms, such as for example algorithms 134 (shown in Figure 3) perform an analysis using a Classification And Regression Tree ("CART") based model, which results in a grouping of UW assets by Collateral Usage and Market Value ("CUMV") groups, using Previous Appraisal Amount ("PAA") as the driving variable.

[0101] Two approaches to assess the performance of a CART based model are outlined below. One approach utilizes a ratio of the sum of squared error (SSE) of a CART based approach to that of a simple model, called an error ratio. A simple model is a model which assigns an average asset price to all assets. The second approach computes a coefficient of determination, denoted as R^2 , and defined as:

[0102] $R^2 = 1 - (SSE/SST)$, where SST is a sum of squares total.

[0103]

R^2 is the contribution of a single asset within each segment relative to the entire population, a higher R^2 value for an asset within a particular segment, the higher is the contribution. The different portfolio segments are ranked based on the two approaches giving an indication of how good the predictive capabilities of the model are within each portfolio segment, giving a comfort level to the bidder in terms of pricing, for example, each tranche

[t4]

(UPB) amounts, region or customer risk. Table C above is an example of defined segments based on tranches and asset rankings (B or C).

[0105] Table C provides an example output from a study of a portfolio with five tranches and two different asset types (B and C). The table shows how the error ratio is ranked for the different segments. Also, the R^2 values for each asset are also computed for assets of type C within each segment.

[0106] A second step is to compute SSE values for each portfolio segment of interest for the CART model and for the simple model (extrapolation of an average price). An error ratio is computed from the SSE based on the CART model divided by an SSE based on the simple model. If the error ratio is less than one, then the CART based model is a better predictor than the simple model. As an added benefit, a superior model can be assembled as a "hybrid" combination of the CART and simple models, by choosing the model which performs best in each segment, according to the error ratio metric.

[0107] A third step is to compute R^2 values for each asset within each portfolio segment. R^2 per asset is computed as $(SST \text{ per segment} - SSE \text{ per segment}) / (\text{overall SST for all assets} \times \text{number of assets within each segment})$.

[0108] Lastly all the segments are ranked based on the error ratio computed in the second step and the R^2 values computed in the third step. The model is accurate in predicting price values for segments that rank high on both of the two metrics, the error ratio and R^2 and superior models are assembled using these metrics.

[0109]

Table D shows the relative ranking of the five tranches for the assets of type C (from Table C) on the basis of the two performance metrics.

[t1]

Table 12 Portfolio Segment Ranking

Tranche CO	C	R-Squared	Rank Error Ratio	Rank R-squared
CO.01	0.73	0.18%	2	2
CO.02	0.61	0.06%	1	5
CO.03	1.58	0.46%	5	1
CO.04	1.47	0.11%	4	4
CO.05	1.20	0.14%	3	3

[0110]

Figure 10 is a flow diagram illustrating an exemplary embodiment of forming clusters 246 using FCM to choose clusters for modeling. Computer 38 (shown in

Figure 2) forms clusters 246 by taking selected data 78 and performing FCM analysis to produce the clusters.

[0111] Figure 11 illustrates building models 248, selecting best models 250 and calculating counts 252 in which six models are built using database 76. Computer 38 (shown in Figure 3) performs this process. Model building 248 is used to assist the underwriter in prioritizing assets for full underwriting 14 and sample-based underwriting 34, as well as for inferential valuation.

[0112]

Figure 11 illustrates building models 248, selecting best models 250 and calculating counts 252 in which six models are built using database 76. Computer 38 (shown in Figure 3) performs this process. Model building 248 is used to assist the underwriter in prioritizing assets for full underwriting 14 and sample-based underwriting 34, as well as for inferential valuation.

The lower portion of Figure 11 is a table illustrating an exemplary embodiment of selecting best models 250 from six models built in accordance with building models 248d. The models differ according to which variables are used as X's. All models use CUMV Cluste (these are present for all assets). The models from building models 248 are used to predict Court Auction Value ("CAV") 256 in addition to Market Value ("MAV") 258. Other embodimen (not shown) use other models to predict other values.

In selecting best models 250, the best models of K regression models under consideration (here, K = 6), are selected. The best model is chosen for each UW asset, according to the following metric $\min_k \{abs(y - \hat{y}_k), 1E-8\}$, where y is the UW value to be predicted, and \hat{y}_k is a prediction from the kth regression model, for k = 1, 2, . . . K.

In calculating counts 252 the number of times each of the K models is selected within each CUMV cluster is counted. Figure 11 contains these counts for CAV and MAV modeling scenarios. Other modeling scenarios are used in other embodiments.

When applying models 254, the weighted average prediction from all models that yielded a prediction for each non-UW asset is used. The weights are constructed from the frequencies of the counts calculated 252, and the predictions come from the modeling process. In one embodiment, a commercial statistical analysis software (SAS) system is used to produce the models. An artifact of using the SAS system is that each non-UW asset will get a predicted UW value from each model for which the non-UW asset has each input variable, i.e., "X variable" present. Other modeling packages share this trait.) Equation E below details the procedure.

$$\hat{y}_i = \frac{\sum_{k=1}^K f_{ik} \hat{y}_k}{\sum_{k=1}^K f_{ik}} \quad (\text{Equation E})$$

In Equation E, f_{ik} = 1 if model k produced a prediction for asset i, and is zero otherwise, f_{ik} = count of times model k was selected for UW assets among the ith CUMV type (i = 1,2), and the jth CUMV cluster (j = 1,2,3), and \hat{y}_k = prediction for y, from model k. Note there is only a contribution from each modeling approach for which an asset has a prediction, with each being weighted by the number of times the modeling approach was selected for all UW assets of the same CUMV cluster.

Process 240 is also used to estimate a Lower Confidence Limit ("LCI") and Upper Confidence Limit ("UCL") for the mean prediction, with a substitution of the corresponding statistic for \hat{y}_k in Equation E

[0113]

Referring back again to Figure 3, supervised learning process 206 and unsupervised learning process 208 use clustering. "Clustering" is a tool that

attempts to assess the relationships among patterns of the data set by organizing the patterns into groups or clusters such that patterns within a cluster are more similar to each other than are patterns belonging to different clusters. That is, the purpose of clustering is to distill natural groupings of data from a large data set, producing a concise representation of a system's behavior. Unsupervised learning step 208, employs a fuzzy clustering method ("FCM") and knowledge engineering to group assets automatically for valuation. FCM is a known method that has been widely used and applied in statistical modeling. The method aims at minimizing intra-cluster distance and maximizing inter-cluster distance. Typically the Euclidean distance is used.

[0114]

FCM 248 (see Figure 10) at the same time minimizes the intra-cluster distance and maximizes the inter-cluster distance. Typically the Euclidean distance is used. FCM is an iterative optimization algorithm that minimizes the cost function

$$J = \sum_{i=1}^n \sum_{k=1}^c \mu_{ik}^m \|X_k - V_i\|^2 \quad (\text{Equation F})$$

where n is the number of data points, c is the number of clusters, X_k is the k^{th} data point, V_i is the i^{th} cluster centroid, μ_{ik} is the degree of membership of the k^{th} data in the i^{th} cluster, m is a constant greater than 1 (typically $m=2$). Note that μ_{ik} is a real number and bounded in $[0,1]$. $\mu_{ik}=1$ means that k^{th} data is definitely in i^{th} cluster, while $\mu_{ik}=0$ means that k^{th} data is definitely not in i^{th} cluster. If $\mu_{ik}=0.5$, then it means that k^{th} data is partially in i^{th} cluster to the degree 0.5. Intuitively, the cost function would be minimized if each data point belongs exactly to a specific cluster and there is no partial degree of membership to any other clusters. That is, there is no ambiguity in assigning each data point to the cluster to which it belongs.

The degree of membership μ_{ik} is defined by

$$\mu_{ik} = \frac{1}{\sum_{j=1}^c \left(\frac{\|X_k - V_i\|^2}{\|X_k - V_j\|^2} \right)^{\frac{1}{m-1}}} \quad (\text{Equation G})$$

Intuitively, μ_{ik} , the degree of membership of the data point X_k in the cluster centroid V_i , increases as X_k is getting closer to V_i . At the same time, μ_{ik} would get smaller as X_k is getting farther away V_j (other clusters).

The i^{th} cluster centroid V_i is defined by

$$V_i = \frac{\sum_{k=1}^n (\mu_{ik})^m X_k}{\sum_{k=1}^n (\mu_{ik})^m} \quad (\text{Equation H})$$

Intuitively, V_i , the i^{th} cluster centroid, is the weighted sum of the coordinates of X_k , where k is the number of data points

Starting with a desired number of clusters c and an initial estimate for each cluster center V_i , $i=1,2,\dots,c$, FCM will converge to a solution for V_i that represents either a local minimum or a saddle point of the cost function. The quality of the FCM solution, like that of most nonlinear optimization problems, depends strongly on the choice of initial values—the number c and the initial cluster centroids V_i

11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

[0115]

In one exemplary embodiment, the entire portfolio 12 is segmented by unsupervised fuzzy clustering and each cluster is reviewed by under-writing

experts, thereby assisting the underwriters in choosing the financial instruments for full underwriting 14 and sample underwriting 34. Alternatively, this FCM can be applied just to portion 42. As a result, each cluster gets assigned a HELTR composite score for purposes of adjustment 138 (see Figure 3). In essence the HELTR composite score captures both expected and range of cash flow, its timing and the risk associated with each cluster.

[0116] Referring now to Figure 2, the ratio of full underwrite portion 16 to the total portfolio 12 is in one exemplary embodiment 25% of the assets and 60% of the face value of all assets. Full underwriting of these assets is warranted due to their size and value. However, this underwriting is fairly uniform for all underwriters, so the underwriting is not likely to produce significant bidding variances. The remaining 40%, however, comprising portions 36 and 42, which in the exemplary embodiment constitute 75% of the assets but only 40% of the face value are highly speculative until underwritten. To the extent value can be found in portions 36 and 42f, for example without limitation, an additional five percent over gross extrapolation, the difference meaning the difference between winning and losing the entire portfolio bid or the entire tranche bid meaning hundreds of millions of dollars difference in profit.

[0117] In the case of insurance policies, in accordance with procedure 40, statistics are used in an attempt to answer three basic questions: (a) How should we collect our data?, (b) How should we summarize the data we collected?, and (c) How accurate are our data summaries?. Algorithm 134 answers question (c) and is a computer-based method without complicated theoretical proofs. Algorithm 134 for insurance policy inferential valuations is suitable for answering statistical inferences that are too complicated for traditional statistical analysis. Algorithm 134 for insurance policy valuation simulates the distribution of statistical estimates by repeatedly sampling with replacement. The algorithm generally is composed of three main steps: (I) Sampling with replacement, (II) Evaluating statistics of interest, and (III) Estimating standard deviation.

[0118] In accordance with insurance algorithm 134, estimates of NPV standard error are

11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

performed as follows. For each of the risk models and for each segment in the models, assuming there are N policies in the segment, n samples are selected using sampling with replacement (for example, $n=100$). Each sample contains N policies, too, in this example. For each sample, and for all historical policies:

$$\frac{A}{E} = \frac{\sum (Act)}{\sum (Wtexp)} \quad (\text{Equation I})$$

Next, net present value is generated by $NPV = \sum P - \sum E - (\sum C) \cdot \frac{A}{E_w}$ (Equation J)

for recent policies. Compute the sample standard deviation for the n NPV values. In Equation I, Act is the actual claim and Wtexp is the weighted expected claim for each individual policy.

[0119]

Figure 12 is a table of exemplary criteria 80 and exemplary rule sets for credit scoring 138. Other criteria could be selected depending on the type of financial

instrument and particular bidding conditions or any other desires or preferences of the bidder.

[0120] Figure 13 is a more detailed tree chart diagram 260 similar to tree chart 66 (see lower portion of Figure 2). In Figure 13, the segregation is by (a) whether secured, (b) whether revolving, (c) whether the last payment was zero. The result is six clusters 262, 264, 266, 268 270, 272, casually known as a "shaker tree".

[0121] Figure 14 illustrates an exemplary system 300 in accordance with one embodiment of the present invention. System 300 includes at least one computer configured as a server 302 and a plurality of other computers 304 coupled to server 302 to form a network. In one embodiment, computers 304 are client systems including a web browser, and server 302 is accessible to computers 304 via the Internet. In addition, server 302 is a computer. Computers 304 are interconnected to the Internet through many interfaces including a network, such as a local area network (LAN) or a wide area network (WAN), dial-in-connections, cable modems and special high-speed ISDN lines. Computers 304 could be any device capable of interconnecting to the Internet including a web-based phone or other web-based connectable equipment, including wireless web and satellite. Server 302 includes a database server 306 connected to a centralized database 76 (also shown in Figure 2) which contains data describing sets of asset portfolios. In one embodiment, centralized database 76 is stored on database server 306 and is accessed by users at one of computers 304 by logging onto server sub-system 302 through one of computers 304. In an alternative embodiment centralized database 76 is stored remotely from server 302. Server 302 is further configured to receive and store information for the asset valuation methods described above.

[0122] While system 300 is described as a networked system, it is contemplated that the methods and algorithms described herein for examination and manipulation of asset portfolios are capable of being implemented in a stand-alone computer system that is not networked to other computers.

[0123] Figure 15 is a diagram 320 illustrating due diligence tools and processes implemented within system 300. Diagram 320 includes repositories of due

diligence information including a dictionary 322, a tool library 324, a deal pitch repository 326, links 328 and dashboards 330 where overviews regarding due diligence information can be accessed. A workspace 332 allows storage of due diligence information being compiled by a due diligence team. A results area 334 includes previously stored results of prior exercises including bid results. A deal valuation area 336 includes stored valuations of former and current deals for comparison.

[0124] The system informally referred to as a Virtual War Room diagrammed in diagram 320 includes a high level map and associated descriptions of the due diligence roles and responsibilities such that collaborators can see who has functional responsibilities, how the team members as individuals fit into the due diligence exercise and who to contact for assorted information. Also included is a project timeline with milestones and tasks arranged as both Gantt charts, PERT charts and text such that key deliverable timing is developed with inputs from the team, and then is made available to a global due diligence team. The project timeline serves as a control mechanism to keep the due diligence on schedule and to account for what if changes to schedule.

[0125] A project feedback mechanism within dashboard 330 includes an easily readable set of graphical indicators which tracks key due diligence deliverables, for example, types and quantities of underwriting completed, total project budget, and status of deliverables. A project calendar with notable local and global dates, holidays, vacations, and deliverables identified is within the feedback mechanism as is contact information for team members and collaborator's including telephone, email and address information. A project to do list and status such that the global team captures vital tasks to be completed and keeps visibility on the status thereof is further contemplated.

[0126] System 300 (shown in Figure 14) includes a shared storage place for the various functions to keep the project files and information in such a fashion that the collaborators can easily have access to the information with ease of retrieval. A centralized information flow map that identifies the sources and uses of

information utilized to make the decisions needed on the due diligence team objectives. An example is the conversion of raw data, through associated calculations and data manipulations in order to set the bid price.

[0127] A centralized repository of system 300 stores the financial models and data manipulation and business process tools that are available for use on the due diligence. The centralized repository for the historical best practices that the business collates and codifies from past due diligence exercises, or elsewhere such that the team can easily deploy and recall the firm's knowledge and lessons learned from past experiences. A centralized gold standard database of relevant valuation information and facts associated with the due diligence. This database is checked for data integrity and has limited access with controlled backups. Underwriting data is centrally stored as is the valuation process so that all underwriting is not predicated by an onsite presence, but can be effected remotely.

[0128] While the invention has been described in terms of various specific embodiments, those skilled in the art will recognize that the invention can be practiced with modification within the spirit and scope of the claims.